Remote Heart Rate Determination in RGB Data
An Investigation using Independent Component Analysis and Adaptive Filtering

Christian Wiede, Julia Richter, André Apitzsch, Fajer KhairAldin and Gangolf Hirtz
Department of Electrical Engineering and Information Technology,
Chemnitz University of Technology, Reichenhainer Str. 70, 09126 Chemnitz, Germany

Keywords: Heart Rate Detection, rPPG, Vital Parameters, Image Processing, Independent Component Analysis, Adaptive Filtering.

Abstract: An emerging topic in the field of elderly care is the determination and tracking of vital parameters, such as the heart rate. This parameter provides important information about a person's current health status. Within the last years, various research focussed on this topic. The recognition of vital parameters is increasingly relevant for our ageing society. This paper presents a method to remotely determine the human heart rate with a camera. At this point, we suggest to use independent component analysis (ICA) and adaptive filtering for a robust detection. In our processing chain, we used different image processing techniques, e. g. face detection, and signal processing techniques, e. g. FFT and bandpass filtering, in this study. An evaluation with several probands, illuminations, frame rates and different heart rate levels showed that we could achieve a mean error of 4.36 BPM, which corresponds to CAND of 94.45 %, and a speed of 35 fps.

1 INTRODUCTION

In a steadily ageing society, taking care for elderly plays a major role. In the last years, several technical assistance systems have been developed to help elderly people in their daily activities (Meinel et al., 2015) and to call help in case of emergencies (Wohlrab et al., 2015). These works show that it is possible to detect a fall in the home environment and inform relatives or caregivers. However, these systems only act after the emergency occurred. The goal of our project is to detect a person's current health status and act pre-emptively in case of indications for possible emergencies. Thus far, current assistant systems are not able to perform this. They are not able to decide whether a person is sleeping in an armchair or has suffered a circulatory collapse and is unconscious.

One possibility to overcome this problem is to detect human vital parameters by means of optical sensors. These so-called physiological parameters (e. g. heart rate, breath rate and oxygen saturation) are detectable with normal cameras due to the fact that the spatial and temporal resolution of the cameras increased in the last years. Especially the heart rate detection drew attention in recent research (Poh et al., 2010; van Gastel et al., 2014). Inspired by these works, this paper presents a novel possibility to robustly detect the heart rate. We used ICA and an adaptive filtering to robustly detect the heart rate from remote. Furthermore, we realised a real-time implementation, so that this method is suitable for elderly care applications.

A high benefit of this method is its contact-less working mode. This method proved to be very convenient for patients, because they do not have to wear any devices. In that way, effects such as skin irritations and discomfort can be avoided. Measuring the heart rate allows the detection of bradycardia and tachycardia at a very early stage, so that emergencies can be avoided. Besides helping elderly people, such a system would be most beneficial for detecting the health status of neonatals to avoid sudden infant death syndrome, for monitoring a driver's well-being or for triage in hospitals.

This paper is organised as follows. Firstly, there will be an overview of related studies in this field. Secondly, we introduce our method for remote heart rate detection. Thereupon, we present our results, which is followed by a discussion. Finally, we summarise our findings and give an outlook on further developments.

Wiede, C., Richter, J., Apitzsch, A., KhairAldin, F. and Hirtz, G.
Remote Heart Rate Determination in RGB Data - An Investigation using Independent Component Analysis and Adaptive Filtering.
DOI: 10.5220/0005694002400246
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 240-246
ISBN: 978-989-758-173-1

2 RELATED WORK

Measuring the human heart rate is an active field of research in the last century. The oldest and most used approach is the electrocardiography (ECG) developed by Willem Einthoven in 1901. An electrocardiograph measures the electric potential between two points on the skin of a human body. The resulting characteristic curves originates from the conduction system of the heart. By simply counting the existing spikes, it is possible to determine the human heart rate. It is the gold standard even nowadays.

Another method evaluates the volumetric changes of the tissue which are caused by the blood flow. This method is called plethysmography. The heart rate can be measured by the variations of air pressure, impedance or strain. In 1937, Hertzman and Spealman recognised the potential of an optical method that later was called photoplethysmography (PPG) (Hertzman and Spealman, 1937). When light is transmitted through a tissue, it changes its wavelength (modulation) depending on the blood flow. The crucial part is the pulsatile fraction of the arterial blood flow, which is only one percent of the overall transmitted light. With this method, thin body parts, such as fingers or earlobes, are penetrated by light emitted by a photo diode (Allen, 2007). The light that was not absorbed is measured on the other side of the body part. This method is also called transmissive PPG, and it is often used for pulse oximetry. Next to the transmissive PPG there is the reflectance PPG that measures the reflected light that is emitted from the tissue. The SNR of the reflective PPG is one dimension smaller than the SNR of the transmissive PPG. However, the reflective PPG has the disadvantage of being contact-based. Moreover, it requires external light sources.

An alternative way is the remote photoplethysmography (rPPG), which is contact-less and does not need any external light source. The basic concept goes back to the year 2005 and was described by Humphrey et al. (K. Humphreys and Ward, 2005). In 2007, Garbey et al. showed a possibility to solve this challenge with a thermal camera (Garbey et al., 2007). The first idea to measure the human heart rate with rPPG in the visible light spectrum was published by Verkruysse et al. in 2008 (Verkruysse et al., 2008). They recorded videos of the human face with a RGB camera and did not use any other light sources about from daylight and normal artificial light. This ambient light could be considered as an additional source of noise. The distance between camera and propositi was 1-2 m. The propositi were instructed not to move while sitting to avoid movement artefacts. A ROI was selected in the persons' faces and a spatial averaging was applied to all three colour channels. They determined the heart rate in an image sequence by using Fast Fourier Transform (FFT). The heart rate is represented by the change of illumination in face.

Although these publications have a basic functionality, they suffered from a low accuracy and artefacts in the signal (e. g. moving artefacts). Poh et al. proposed a more advanced method (Poh et al., 2010; Poh et al., 2011). They used independent component analysis (ICA) as a blind source separation for the three colour channels. This resulted in a more stable heart rate determination that is robust against small motion artefacts. Later, this general idea was improved by van Gastel et al. (van Gastel et al., 2014) by using only the forehead region and by applying several temporal filters. Instead of ICA, Lewandowska et al. (Lewandowska et al., 2011) suggested to use principal component analysis (PCA), which gives accurate results while less computational power is needed. In another publication, chrominance-based rPPG was introduced. At this point, colour difference signals were used to eliminate specular reflections on the skin (de Haan and Jeanne, 2013). A different way to solve the problem of rPPG is the Eulerian video magnification (Wu et al., 2012; Rubinstein, 2014). With this method, small movements in images can be visualised to the human eye. He et al. showed how to detect rPPG with this method (He et al., 2014). An alternative way that uses Newtonian reaction is presented by Balakrishnan 2013 (Balakrishnan et al., 2013). They detected certain interest points in the images, applied PCA and determined small movements caused by the cyclic blood movement. Other methods demonstrated the possibility to make rPPG invariant against motion as well (Li et al., 2014; Wang et al., 2015).

In 2015, van Gastel et al. proved that rPPG also works in the near infrared spectrum (van Gastel et al., 2015). They showed that the same methods are applicable for these wavelengths.

On the basis of pulse detection, it is possible to extract other important medical parameters such as the morphology of the signal or the heart rate variability (HRV).

In our paper, we propose an approach to combine ICA and adaptive filtering for a robust heart rate determination.

3 METHODS

3.1 System Overview

The following section describes the methods we applied in order to design a system for rPPG. Figure 1 shows the major steps. At first, images are acquired.

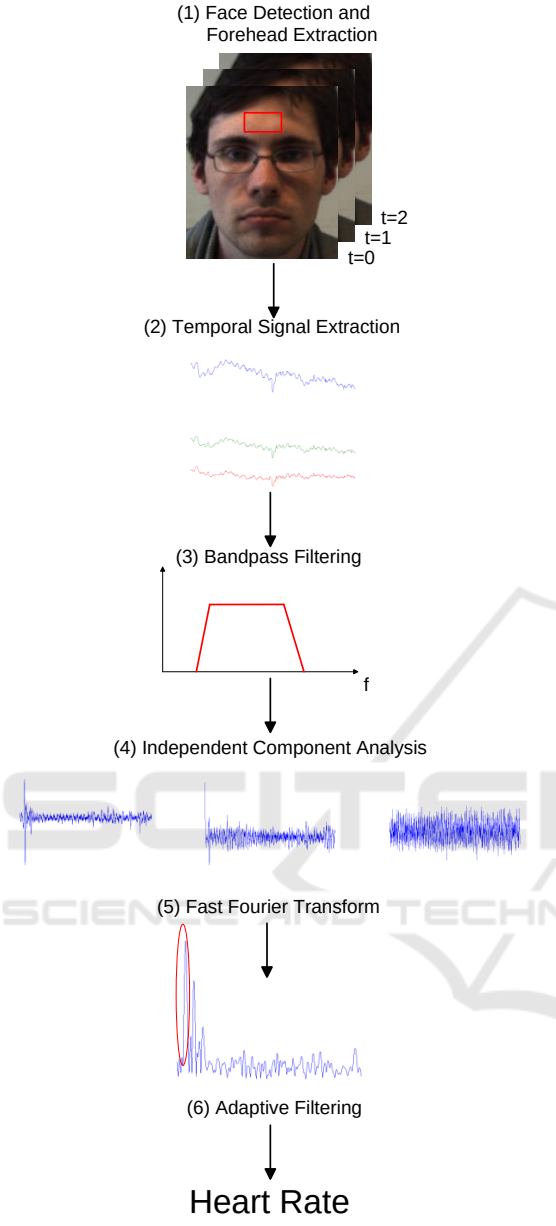

Figure 1: This overview shows the system functionality. First of all, the face is detected and a forehead region is defined in the image (1). Then, three mean temporal signals are extracted from the RGB channels (2). This step is followed by a bandpass filter (3). The ICA splits the three colour channels into three independent components (4). By using Fast Fourier Transform, existing frequencies in these components are determined and the highest peak in the spectrum is detected as the preliminary heart rate (5). With an adaptive filtering, quick changes in the signal can be compensated (6), so that as final result, the heart rate can be obtained.

This step is followed by a face detection, the selection of the forehead region and the extraction of the three colour channels. A bandpass limits the frequencies to the natural limits of the human heart rate. Afterwards, an ICA is applied to the three channels. A frequency analysis determines the final heart rate. With the aid of an adaptive filtering, outliers can be eliminated.

A normal RGB camera (Basler ac640-100gc) was used for image acquisition. It provides a VGA resolution and alterable frame rates can be adjusted. For each recording, the frame rate was set to a fixed value (30 fps or 50 fps) to have equidistant time steps for the signal processing. The automatic controls for exposure time and white balancing were switched off to avoid invalid changes in the signal.

3.2 Face Detection and Forehead Region Extraction

Due to the fact that our method works colour-based, we need regions in the image where skin pixels are visible. At this point, the face is the most suitable region. Normally, the face is not occluded by garments and usually has a vertical orientation, which simplifies the detection.

In the last decades, face detection is a well-studied topic. We used the Viola-Jones Detector (Viola and Jones, 2004) in OpenCV, which classifies Haar-like features in cascades. The face detection itself generates bounding boxes with the face inside. There are several face regions that are not suitable for heart rate determination, because they do not provide significant information, e. g. hair and eyebrows, or because they show strong movement artefacts, e. g. eyes and mouth. Therefore, only a part of the face region was taken into account. For a better signal quality, a region of interest (ROI) was selected that provides a relatively low noise-affected heart rate signal. For this, the forehead region has been chosen, since it disposes sufficient superficial vessels because of the thin skin. Moreover, it shows a good light reflection characteristic with a low light absorption of the tissue. Furthermore, movement artefacts are significantly less frequent than in other face regions. The ROI is defined by a static window within the face region that has always the same position with respect to bounding box of the face.

3.3 Temporal Signal Extraction and Band Pass Filtering

After having selected the forehead region, all pixels inside the ROI are summarised and averaged. These

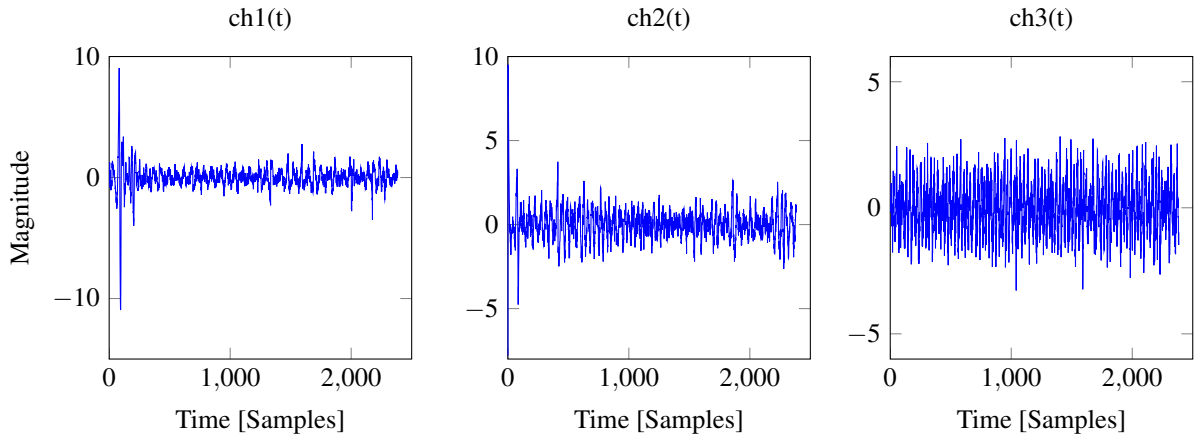

Figure 2: In result of the ICA, there are three independent components. The third component ch3(t) showed the best result in our implementation.

operations are performed for all three RGB colour channels:

$$R_{mean}(t_0) = \frac{1}{n_{roi}} \sum_{i}^{roi} \sum_{j}^{roi} R_{i,j}(t_0) \tag{1a}$$

$$G_{mean}(t_0) = \frac{1}{n_{roi}} \sum_{i}^{roi} \sum_{j}^{roi} G_{i,j}(t_0) \tag{1b}$$

$$B_{mean}(t_0) = \frac{1}{n_{roi}} \sum_{i}^{roi} \sum_{j}^{roi} B_{i,j}(t_0) \tag{1c}$$

$R_{i,j}$, $G_{i,j}$ and $B_{i,j}$ denote a single pixel in the ROI of the corresponding channel. n_{ROI} is the number of all pixels in the ROI. Consequently, for every specific moment in time t_0, there is one value for each colour channel that represents the mean value R_{mean}, G_{mean} and B_{mean}.

In order to exclude implausible frequencies that cannot represent the human heart rate, a bandpass filter BP is applied, see Equation (2a) to Equation (2c). Only frequencies higher than 0.5 Hz (30 BPM) and lower than 3 Hz (180 BPM) are considered during further computations, i. e. the signals R_{BP}, G_{BP} and B_{BP}. For our practical implementation, we used a FIR filter with 201 filter components. This filter has a linear phase response, which means that all frequencies have the same group delay. Moreover, the filtering has the effect of a pre-whitening, which is necessary for the ICA.

$$R_{BP}(t) = BP(t) * R_{mean}(t) \tag{2a}$$

$$G_{BP}(t) = BP(t) * G_{mean}(t) \tag{2b}$$

$$B_{BP}(t) = BP(t) * B_{mean}(t) \tag{2c}$$

3.4 Independent Component Analysis

The three colour channels contain several sources of image noise and artefacts, e. g. motion. The objective is to find the underlying, original signals and extract the pulse signal by decomposing the colour channels. One possibility for decomposition is the ICA. This method assumes that our observations are a linear combination of the independent sources. Hence, it is called blind source separation. In the general equation $\vec{x} = \mathbf{A}\vec{s}$, $\vec{x}$ denotes the vector of the observed components, $\mathbf{A}$ is the so-called mixing matrix with linear concatenated elements and $\vec{s}$ represents the independent source components. For this application, the equation can be formulated as:

$$\begin{bmatrix} R_{BP}(t) \\ G_{BP}(t) \\ B_{BP}(t) \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \begin{bmatrix} ch_1(t) \\ ch_2(t) \\ ch_3(t) \end{bmatrix} \tag{3}$$

For this implementation, we used the FastICA approach of Hyvärinen (Hyvärinen, 1999). It is accurate, fast and available for several programming languages.

In Figure 2, all three components are displayed. The component with the highest periodicity in the signal, which is visible in the harmonics of the spectrum, is most likely the component we are looking for. In our study, the third independent components $ch_3(t)$ shows the highest periodicity. In order to remove the high-frequent noise that was caused by the ICA, we applied a smoothing window, which averages over three samples. This lowpass is denoted as LP.

$$ch_{3smooth} = LP(t) * ch_3(t) \tag{4}$$

The smoothed result $ch_{3smooth}$ is shown in Figure 3.

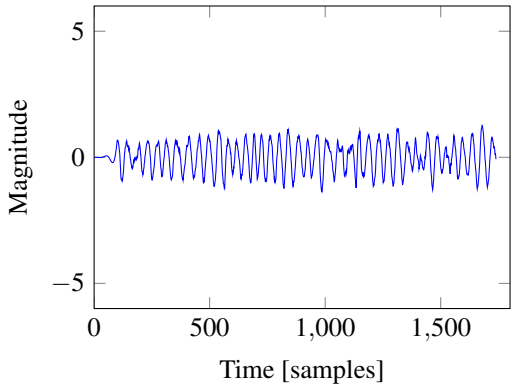

Figure 3: This plot shows the smoothed third independent component.

3.5 Frequency Analysis

Finally, the frequencies of the signal $ch_{3smooth}$ are determined using the Fast Fourier Transform (FFT). Since the heart rate can change over time, we performed the FFT in a small window of three seconds. In this short interval, the human heart rate will not change considerably. In frequency domain, the most prominent peak is selected. This peak represents the heart rate HR at time t_0.

$$\mathrm{HR}(t_0) = \max(|\mathrm{FFT}(ch_{3smooth})|) \tag{5}$$

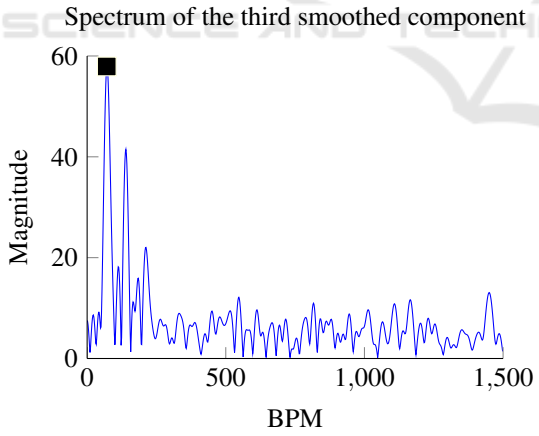

Figure 4: The Fast Fourier Transform provides the frequency spectrum of the third smoothed component. To increase the frequency resolution, zero padding was applied. The maximum peak is for this example at 72.5 BPM.

In order to increase the frequency resolution, we applied zero padding to the signal. This guarantees a more accurate quantisation without changing the signal information.

3.6 Adaptive Filtering

In order to eliminate false detections appearing in the form of sudden changes in the heart rate from one window to the next, we suggest the usage of an adaptive filter. If the absolute difference between the current heart rate value $\mathrm{HR}(t_0)$ and the previous heart rate value $\mathrm{HR}(t_0 - \Delta t)$ is higher than a threshold T) of 10 BPM, a sliding average is computed using the last 10 values according to Equation (6) and Equation (7). Δt denotes the time interval between to samples.

$$\mathrm{T} = |\mathrm{HR}(t_0 - \Delta t) - \mathrm{HR}(t_0)| \tag{6}$$

$$\mathrm{HR}_{ad} = \begin{cases} \mathrm{HR}(t_0) & \text{if T} < 10 \text{ BPM} \\ \frac{1}{10}\sum_{t=t_0-9}^{t_0} \mathrm{HR}(t) & \text{otherwise.} \end{cases} \tag{7}$$

4 EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Setting

For our experiments, we used a Basler RGB camera (acA640-100gc). This camera records data with a fixed frame rate. Both automatic exposure time control and automatic white balancing are switched off. As far as lighting sources are concerned, we tested multiple subjects under different lighting conditions, whereas the only lighting sources were daylight and the illumination of the interior lights in the room. The distance between camera and the probands was between one and two meters. The probands were instructed to sit still with the face directed to the camera. The recorded sequences have a length between 30 seconds and 2 minutes.

As reference we used a Polar FT1 heart rate monitor. This system is composed of a heart rate sensor, which is mounted on a chest strap, and a display, which shows the current heart rate. The display is visible in the corresponding videos as well, with the result that the current reference heart rate can be compared with our results.

In order to generate more variability in the heart rate, the probands were asked to perform squats before the recording of some sequences. This results in higher heart rates and is more representative with respect to the natural range of the human heart rate.

In the next sections the accuracy, the utilization of other colour channels than RGB and the computational speed are considered and discussed.

4.2 Accuracy

In our recordings, the average heart rate ranged from 67 BPM to 114 BPM. The mean error is calculated for each window from t_0 to t_{end} as the absolute difference between the measured heart rate $\mathrm{HR}_{ad}(t)$ and the reference heart rate $\mathrm{HR}_{Ref}(t)$, see equation 8. To measure the accuracy, we used this mean error and the complement of the absolute normalised difference (CAND) as evaluation criterion, as shown in equation 9.

$$\mathrm{Mean} = \sum_{t=t_0}^{t_{end}} \left|\mathrm{HR}_{Ref}(t) - \mathrm{HR}_{ad}(t)\right| \tag{8}$$

$$\mathrm{CAND} = 1 - \frac{\left|\mathrm{HR}_{Ref} - \mathrm{HR}_{ad}\right|}{\mathrm{HR}_{Ref}} \tag{9}$$

In our publication, the CAND is expressed in percent. The higher this value, the better is the algorithm. In our overall results, the mean error of the heart rate detection is 4.36 BPM and the CAND is 94.45 %. This shows that we can accurately detect the human heart rate with our method. For our addressed application fields, these errors are reasonably small.

Table 1: Mean error and the CAND for the different image sequences.

Sequence	Mean Error in BPM	CAND in %
Sequence 1	5.63	92.34
Sequence 2	2.59	96.27
Sequence 3	3.36	95.24
Sequence 4	5.54	94.77
Sequence 5	4.85	93.64

We demonstrated that the ICA and the adaptive filtering reduce the error significantly. However, the problem of ICA is the selection of the most suitable component. Methods such as autocorrelation or spectral density could solve this issue.

In another experiment, we tested our algorithm for another colour space, the HSI. The HSI color space splits the channels in hue, saturation and intensity. We expected that the decoupling of intensity and hue shows better result than RGB. However, the results were a tenfold worse. That indicates that our proposed method is not suitable for the HSI colour space.

4.3 Computational Speed

In order to increase the speed of our algorithm, we reimplemented the Matlab algorithm in C++ by using the computer vision library OpenCV. For the C++ implementation, we measured the computational speed. For the tests, an Intel core i7 quad core processor with 2.9 GHz and 16 GB RAM was used. For compiling GCC 5.2 was used.

In the case of computing the ICA in every time step, the mean processing time, starting from image acquisition to the derivation of the current heart rate, is 28 ms. This corresponds approximately to 35 fps, which demonstrates that our method runs fluently on the described system. A speed comparison with other publications was not possible, because this information is not provided in literature.

A possibility to increase the speed is to calculate the mixing matrix of the ICA only every n windows or using the result of the ICA of the previous time step for initialisation. Another acceleration option is to use only R and G components of the RGB channel for the ICA, because we assume that these two channels contain the main part of the actual heart rate signal. As a consequence, the computational effort will be less with probably the same accuracy.

5 CONCLUSIONS

In this paper, we presented a method for remote heart rate determination using ICA and adaptive filtering. With regard to applications in domestic environments and for elderly care, the obtained results are adequate. For other use cases, especially in clinical environments, where highly accurate measurements are required, accuracy has to be improved.

For future work, the algorithm should also be robust against motion artefacts. One solution could be a feature tracking on the forehead region.

By porting the algorithm to an embedded system, a more flexible and praxis-oriented solution could be achieved.

With the help of such a system, it could be possible to pre-emptively detect emergencies in domestic environments.

ACKNOWLEDGEMENTS

This project is funded by the European Social Fund (ESF). We would like to thank all probands who took part in the experiments and supported us with their video records.

REFERENCES

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):R1–R39.

Balakrishnan, G., Durand, F., and Guttag, J. (2013). Detecting pulse from head motions in video. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3430–3437.

de Haan, G. and Jeanne, V. (2013). Robust pulse rate from chrominance-based rppg. *Biomedical Engineering, IEEE Transactions on*, 60(10):2878–2886.

Garbey, M., Sun, N., Merla, A., and Pavlidis, I. (2007). Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *Biomedical Engineering, IEEE Transactions on*, 54(8):1418–1426.

He, X., Goubran, R., and Liu, X. (2014). Using eulerian video magnification framework to measure pulse transit time. In *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on*, pages 1–4.

Hertzman, A. B. and Spealman, C. R. (1937). Observations on the finger volume pulse recorded photoelectrically. *American Journal of Physiology*, 119:334–335.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634.

K. Humphreys, C. M. and Ward, T. (2005). A cmos camera-based system for clinical photoplethysmographic applications. In *Proceedings of SPIE*, volume 5823, pages 88–95.

Lewandowska, M., Ruminski, J., Kocejko, T., and Nowak, J. (2011). Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity. In *Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on*, pages 405–410.

Li, X., Chen, J., Zhao, G., and Pietikainen, M. (2014). Remote heart rate measurement from face videos under realistic situations. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 4264–4271.

Meinel, L., Richter, J., Schmidt, R., Findeisen, M., and Hirtz, G. (2015). Opdemiva: An integrated assistance and information system for elderly with dementia. In *Consumer Electronics (ICCE), 2015 IEEE International Conference on*, pages 76–77.

Poh, M.-Z., McDuff, D., and Picard, R. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762–10774.

Poh, M.-Z., McDuff, D., and Picard, R. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. *Biomedical Engineering, IEEE Transactions on*, 58(1):7–11.

Rubinstein, M. (2014). *Analysis and Visualization of Temporal Variations in Video*. PhD thesis, Massachusetts Institute of Technology.

van Gastel, M., Stuijk, S., and de Haan, G. (2015). Motion robust remote-ppg in infrared. *Biomedical Engineering, IEEE Transactions on*, PP(99):1–1.

van Gastel, M., Zinger, S., Kemps, H., and de With, P. (2014). e-health video system for performance analysis in heart revalidation cycling. In *Consumer Electronics Berlin (ICCE-Berlin), 2014 IEEE Fourth International Conference on*, pages 31–35.

Verkruysse, W., Svaasand, L. O., and Nelson, J. S. (2008). Observations on the finger volume pulse recorded photoelectrically. *Optics Express*, 16(26):21434–21445.

Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.

Wang, W., Stuijk, S., and de Haan, G. (2015). Exploiting spatial redundancy of image sensor for motion robust rppg. *Biomedical Engineering, IEEE Transactions on*, 62(2):415–425.

Wohlrab, D., Heß, M., Apitzsch, A., Langklotz, M., Schwarzenberger, A., Bilda, S., Schulz, H., Hirtz, G., and Mehner, J. (2015). Hom-e-call - an enhanced fall detection system based on accelerometer and optical sensors applicable in domestic environment. In Jaffray, D. A., editor, *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*, volume 51 of *IFMBE Proceedings*, pages 1461–1464. Springer International Publishing.

Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., and Freeman, W. T. (2012). Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph. (Proceedings SIGGRAPH 2012)*, 31(4).